METHODS AND APPLICATIONS OF ANALYSIS.
Vol. 8, No. 1, pp. 15–32, March 2001


# OPTIMAL ACCELERATION OF CONVERGENCE*

ARIEH ISERLES† AND XIAOYAN LIU‡

**Abstract.** In this paper we investigate the *quadratic model*

$$\frac{\mathbf{b}^T\mathbf{u}+\mathbf{u}^TC\mathbf{u}}{\mathbf{a}^T\mathbf{u}},$$

which generalises the Shanks transformation and the familiar Padé model, to accelerate the convergence of a sequence $\{u_k\}_{k\in Z_p}$. In the new model,

$$\mathbf{u}^T=[u_0,u_1,\ldots,u_n],$$

$\mathbf{a},\mathbf{b}\in\mathbb{R}^{n+1}$ and $C$ is an $(n+1)\times(n+1)$ Hermitian matrix. (Note that, although $\mathbf{u}$ might be complex, we use the transpose $\mathbf{u}^T$, rather than the adjoint $\mathbf{u}^*$.) Let us assume that the original sequence is obtained from the dynamical system $z\mapsto f(z)$, where $f(z)=\hat{u}+\sum_{k=1}^{\infty}\frac{1}{k!}\beta_k(z-\hat{u})^k$ is analytic about its fixed point $\hat{u}\in\mathbb{C}$. Suppose that $u_0=\hat{u}+\varepsilon$ and $u_k=f^{\circ k}(u_0)$. It is easy to show that $\mathbf{b}=\mathbf{0}$ and $\mathbf{a}=2C\mathbf{1}$. We give an iterative formula for the construction of the matrices $C$. Furthermore, we discuss the rate of convergence, inclusive of the special case when the fixed point is at $\infty$.

**1. The quadratic model.** The starting point for the Shanks transformation is that, given a sequence $\{u_k\}_{k\in\mathbb{Z}^+}$, we construct the function

$$G(z):=u_0+\sum_{k=1}^{\infty}(u_k-u_{k-1})z^k$$

and take the $[N/N]$ Padé approximant to $F$ at $z=1$ as the 'accelerated' limit [1].

Its obvious generalisation is to consider

$$G(z)=\sum_{m=0}^{\infty}\left[\sum_{j=0}^{m}(-1)^{m-j}\alpha_{m,j}u_j\right]z^m$$

and its sections

$$G^{[n]}(z)=\sum_{m=0}^{n}\left[\sum_{j=0}^{m}(-1)^{m-j}\alpha_{m,j}u_j\right]z^m.$$

Having fixed $n$, we stipulate that $G^{[n]}(1)=u_n$. Thus,

$$\sum_{j=0}^{n}\left[\sum_{m=j}^{n}(-1)^{m-j}\alpha_{m,j}\right]u_j=u_n$$

and, since we want the coefficients to be independent of $\{u_k\}_{k\in\mathbb{Z}^+}$, we stipulate $\alpha_{n,n}=1$ and

$$\sum_{m=j}^{n}(-1)^{m-j}\alpha_{m,j}=0,\qquad j=0,1,\ldots,n-1. \tag{1.1}$$

*Received Sept. 27, 1999; revised April 10, 2001.

†Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England (A.Iserles@damtp.cam.ac.uk).

‡Department of Mathematics and Physics, University of La Verne, USA (liux@ULV.EDU).

This is our first model, and we call it the *Padé model*.

Another general scheme for convergence acceleration is by letting the accelerator be

$$\frac{\mathbf{b}^T\mathbf{u}+\mathbf{u}^TC\mathbf{u}}{\mathbf{a}^T\mathbf{u}}, \tag{1.2}$$

where

$$\mathbf{u}^T=[u_0,u_1,\ldots,u_n]$$

$\mathbf{a},\mathbf{b}\in\mathbb{R}^{n+1}$ and $C$ is an $(n+1)\times(n+1)$ Hermitian matrix. Note that, although $\mathbf{u}$ might be complex, we use the transpose $\mathbf{u}^T$, rather than the adjoint $\mathbf{u}^*$. We call this the *quadratic model.*

**2. The quadratic model.** Let us assume that the original sequence is obtained from the dynamical system $z\mapsto f(z)$, where $f(z)=\hat{u}+\sum_{k=1}^{\infty}\frac{1}{k!}\beta_k(z-\hat{u})^k$ is analytic about its fixed point $\hat{u}\in\mathbb{C}$. Suppose that $u_0=\hat{u}+\varepsilon$. Then, expanding into series,

$$\begin{aligned}u_1&=\hat{u}+\beta_1\varepsilon+\tfrac{1}{2}\beta_2\varepsilon^2+O(\varepsilon^3),\\ u_2&=\hat{u}+\beta_1^2\varepsilon+\tfrac{1}{2}(\beta_1+\beta_1^2)\beta_2\varepsilon^2+O(\varepsilon^3).\end{aligned}$$

Generally, we have (except when $\beta_1=1$)

$$u_k=f^{\circ k}(\hat{u}+\varepsilon)=\hat{u}+\beta_1^k\varepsilon+\tfrac{1}{2}\beta_1^{k-1}\frac{1-\beta_1^k}{1-\beta_1}\beta_2\varepsilon^2+O(\varepsilon^3),\qquad k\in\mathbb{Z}^+, \tag{2.1}$$

where $f^{\circ k}$ is the $k$th iterate of the function $f$. Set

$$V(\varepsilon)=\frac{\mathbf{b}^T\mathbf{u}+\mathbf{u}^TC\mathbf{u}}{\mathbf{a}^T\mathbf{u}}.$$

The *order* of convergence acceleration (equivalently, the order of superattractivity at the fixed point) is the integer $p$ such that

$$V(\varepsilon)=\hat{u}+O(\varepsilon^{p+1}).$$

The condition for $p\geq 0$ is that $V(0)=\hat{u}$. Let

$$\mathbf{p}_\ell:=[1,\beta_1^\ell,\beta_1^{2\ell},\ldots,\beta_1^{n\ell}]^T,\qquad \ell\in\mathbb{Z}^+$$

(for convenience, $\mathbf{p}_0=\mathbf{1}$).

The $\varepsilon^0$ condition is

$$\mathbf{b}^T\mathbf{1}\hat{u}+\mathbf{1}^TC\mathbf{1}\hat{u}^2=\mathbf{a}^T\mathbf{1}\hat{u}^2,$$

and we deduce that

$$\mathbf{b}^T\mathbf{1}=0,\qquad \mathbf{1}^TC\mathbf{1}=\mathbf{a}^T\mathbf{1}. \tag{2.2}$$

We turn our attention next to the $\varepsilon$ condition: it is

$$\mathbf{b}^T\mathbf{p}_1+2\mathbf{1}^TC\mathbf{p}_1\hat{u}=\mathbf{a}^T\mathbf{p}_1\hat{u}.$$

We first deduce that $\mathbf{b}^T\mathbf{p}_1 = 0$, and this implies $\mathbf{b} = \mathbf{0}$. Hence, no need to elaborate or mention any further that vector. More interesting consequence of equating powers of $\beta_1$ is

$$\mathbf{a} = 2C\mathbf{1}. \tag{2.3}$$

Comparing with (2.2), we deduce that

$$\mathbf{1}^T C\mathbf{1} = \mathbf{a}^T\mathbf{1} = 0. \tag{2.4}$$

Thus, we need to go further, considering the $\varepsilon^2$ terms, to argue that $p \geq 1$. This gives (for $\beta_1 \neq 0$ – otherwise the analysis is even simpler!)

$$\mathbf{p}_1^T C\mathbf{p}_1 + \hat{u}\frac{\beta_2}{\beta_1(1-\beta_1)}(\mathbf{p}_1 - \mathbf{p}_2)^T C\mathbf{1} = \tfrac{1}{2}\hat{u}\frac{\beta_2}{\beta_1(1-\beta_1)}(\mathbf{p}_1 - \mathbf{p}_2)^T\mathbf{a}$$

and (2.3) means that the $p \geq 1$ condition is just

$$\mathbf{q}^T C\mathbf{q} \equiv 0 \qquad \text{for all} \qquad \mathbf{q} = [1, q, q^2, \ldots, q^n]^T, \quad q \in \mathbb{R}. \tag{2.5}$$

Hence

$$\sum_{k=0}^{n}\sum_{\ell=0}^{n} c_{k,\ell}q^{k+\ell} = 0, \qquad q \in \mathbb{R},$$

and (2.5) can be alternatively expressed as

$$\begin{aligned} \sum_{k=0}^{\ell} c_{k,\ell-k} &= 0, \qquad \ell = 0, 1, \ldots, n; \\ \sum_{k=\ell-n}^{n} c_{k,\ell-k} &= 0, \qquad \ell = n+1, \ldots, 2n. \end{aligned} \tag{2.6}$$

Note, incidentally, that $\mathbf{1}^T C\mathbf{1} = 0$ is a special case of (2.5) when $q = 1$.

Recall from [2] that there exist numbers $\{\alpha_{m,\ell} : \ell = 1, \ldots, m\, ; m = 1, 2, \ldots\}$ such that

$$f^{\circ k}(z) = \hat{u} + \sum_{m=1}^{\infty}\left(\sum_{\ell=1}^{m}\alpha_{m,\ell}\beta_1^{k\ell}\right)\varepsilon^m$$

(we exclude the cases of $\beta_1 = 0$ and of $\beta_1$ being a root of unity). Consequently, and exploiting $\mathbf{1}^T C\mathbf{1} = 0$, we obtain

$$\begin{aligned} \mathbf{u}^T C\mathbf{u} = {} & 2\hat{u}\sum_{m=1}^{\infty}\left(\sum_{\ell=1}^{m}\alpha_{m,\ell}\mathbf{1}^T C\mathbf{p}_\ell\right)\varepsilon^m \\ & + \sum_{m=2}^{\infty}\left(\sum_{k=1}^{m-1}\sum_{\ell=1}^{k}\sum_{j=1}^{m-k}\alpha_{k,\ell}\alpha_{m-k,j}\mathbf{p}_\ell^T C\mathbf{p}_j\right)\varepsilon^m \end{aligned} \tag{2.7}$$

$$\mathbf{1}^T C\mathbf{u} = \sum_{m=1}^{\infty}\left(\sum_{\ell=1}^{m}\alpha_{m,\ell}\mathbf{1}^T C\mathbf{p}_\ell\right)\varepsilon^m. \tag{2.8}$$

Let

$$W(\varepsilon) := \varepsilon^{-1}\left(\mathbf{u}^T C\mathbf{u} - 2\hat{u}\mathbf{1}^T C\mathbf{u}\right).$$

Then a necessary condition for order $p$ is that $W(\varepsilon) = O(\varepsilon^{p+1})$ (recall that both the numerator and the denominator are $O(\varepsilon)$). Clearly, this is also sufficient if the $O(\varepsilon)$ terms therein are nonzero. Comparing (2.7) and (2.8), we have

$$\begin{aligned} W(\varepsilon) &= \sum_{m=1}^{\infty}\left(\sum_{k=1}^{m}\sum_{\ell=1}^{k}\sum_{j=1}^{m+1-k}\alpha_{k,\ell}\alpha_{m+1-k,j}\mathbf{p}_\ell^T C\mathbf{p}_j\right)\varepsilon^m \\ &= \sum_{m=1}^{\infty}\sum_{\ell=1}^{m}\sum_{j=1}^{m+1-\ell}\left(\sum_{k=\ell}^{m+1-j}\alpha_{k,\ell}\alpha_{m+1-k,j}\right)\mathbf{p}_\ell^T C\mathbf{p}_j\varepsilon^m. \end{aligned} \tag{2.9}$$

Moreover,

$$V(\varepsilon) = \frac{\hat{u}\left[\alpha_{1,1}\mathbf{1}^T C\mathbf{p}_1 + \sum_{m=1}^{\infty}\left(\sum_{\ell=1}^{m+1}\alpha_{m+1,\ell}\mathbf{1}^T C\mathbf{p}_\ell\right)\varepsilon^m\right] + \frac{1}{2}W(\varepsilon)}{\alpha_{1,1}\mathbf{1}^T C\mathbf{p}_1 + \sum_{m=1}^{\infty}\left(\sum_{\ell=1}^{m+1}\alpha_{m+1,\ell}\mathbf{1}^T C\mathbf{p}_\ell\right)\varepsilon^m},$$

hence the method is of order at least $p$ if $\mathbf{1}^T C\mathbf{p}_1 \neq 0$. Thus, the order is degraded for all $\beta_1$ if $C\mathbf{1} = \mathbf{0}$ and it might be degraded for particular values of $\beta_1$ if $C$ is singular and $\mathbf{p}_1$ is an eigenvector corresponding to the eigenvalue 0. We impose in the sequel the requirement that $C\mathbf{1} \neq \mathbf{0}$ (i.e. that $C$ is *nondegenerate*), W.L.O.G..

As we have already seen, $p \geq 1$ only if

$$\alpha_{1,1}^2\mathbf{p}_1^T C\mathbf{p}_1 = 0.$$

Consequently, we require $\mathbf{q}^T C\mathbf{q} \equiv 0\ \forall q \in \mathbb{R}$. The necessary condition for $p \geq 2$ is

$$2\alpha_{1,1}\left(\alpha_{2,1}\mathbf{p}_1^T C\mathbf{p}_1 + \alpha_{2,2}\mathbf{p}_1^T C\mathbf{p}_2\right) = 0,$$

hence we deduce that

$$\mathbf{q}_1^T C\mathbf{q}_1 \equiv \mathbf{q}_1^T C\mathbf{q}_2 \equiv 0, \tag{2.10}$$

where

$$\mathbf{q}_j = [1, q^j, q^{2j}, \ldots, q^{nj}]^T,\ j \in \mathbb{Z}^+,\ q \in \mathbb{R}.$$

**3. The General Case.** We consider next $p \geq 3$. The conditions (2.10) are still necessary and the third-order terms are

$$\begin{aligned} &(2\alpha_{1,1}\alpha_{3,1} + \alpha_{2,1}^2)\mathbf{p}_1^T C\mathbf{p}_1 \\ &+ 2(\alpha_{1,1}\alpha_{3,2} + \alpha_{2,1}\alpha_{2,2})\mathbf{p}_1^T C\mathbf{p}_2 \\ &+ 2\alpha_{1,1}\alpha_{3,3}\mathbf{p}_1^T C\mathbf{p}_3 + \alpha_{2,2}^2\mathbf{p}_2^T C\mathbf{p}_2, \end{aligned}$$

since (2.10) implies $\mathbf{q}_2^T C\mathbf{q}_2 = 0$ (note that $\mathbf{q}_2(q) = \mathbf{q}_1(q^2)$). Hence we require just one more equation,

$$\mathbf{q}_1^T C\mathbf{q}_3 = 0. \tag{3.1}$$

Generally, by the similar reasoning, we need

$$\mathbf{q}_1^T C \mathbf{q}_i = 0, \qquad i = 1, 2, \ldots, p, \tag{3.2}$$

for order $p$. We know that the matrix $C$ cannot be $p \times p$, it has to be larger. In the following proposition we will prove that for $N = p(p+3)/2$, the smallest matrix $C$ satisfying (3.2) is $(N+1) \times (N+1)$. A lemma is required first.

LEMMA 3.1. *Let $C$ be a symmetric matrix, and*

$$C = [a_{k,\ell}]_{k,\ell=0,1,\ldots,K}, \qquad a(t,s) = \sum_{k=0}^{K} \sum_{\ell=0}^{K} a_{k,\ell} t^k s^\ell.$$

*If $a(q^j, q^i) = 0$, $1 \le i \le j$, $i$ and $j$ are relatively prime, then $(t^i - s^j) | a(t,s)$.*

*Proof.* Let us consider the coefficients $a_{k,\ell}$ of $a(t,s)$. Since

$$\begin{aligned} 0 = a(q^j, q^i) &= \sum_{k=0}^{K} \sum_{\ell=0}^{K} a_{k,\ell} q^{jk} q^{i\ell} \\ &= \sum (a_{k,\ell} + a_{k-i,\ell+j} + a_{k-2i,\ell+2j} + \ldots + a_{k-ui,\ell+uj}) q^{jk+i\ell}, \end{aligned}$$

every coefficient of $q^{jk+i\ell}$ has to be zero for $0 \le k \le K$, $0 \le \ell \le K$. That means, using the fact that $\gcd(i,j) = 1$ (hence each coefficient appears only in one equation) that

$$a_{k,\ell} + a_{k-i,\ell+j} + a_{k-2i,\ell+2j} + \ldots + a_{k-ui,\ell+uj} = 0,$$

where all subscripts are bounded by 0 and $K$. Notice that there are complex numbers $a_1, a_2, \ldots, a_u$, $b_1, b_2, \ldots, b_u$ such that

$$\begin{aligned} &a_{k,\ell} t^k s^\ell + a_{k-i,\ell+j} t^{k-i} s^{\ell+j} + a_{k-2i,\ell+2j} t^{k-2i} s^{\ell+2j} + \cdots + a_{k-ui,\ell+uj} t^{k-ui} s^{\ell+uj} \\ &= t^{k-ui} s^{\ell+uj} (a_1 t^i s^{-j} + b_1)(a_2 t^i s^{-j} + b_2) \cdots (a_u t^i s^{-j} + b_u) \\ &= t^{k-ui} s^\ell (a_1 t^i + b_1 s^j)(a_2 t^i + b_2 s^j) \cdots (a_u t^i + b_u s^j). \end{aligned}$$

Substituting $t = 1$ and $s = 1$ we get

$$0 = a_{k,\ell} + a_{k-i,\ell+j} + a_{k-2i,\ell+2j} + \cdots + a_{k-ui,\ell+uj} = (a_1 + b_1)(a_2 + b_2) \cdots (a_u + b_u).$$

One of $(a_m + b_m)$s must be zero and, without loss of generality, we assume that it is the first. Hence $a_1 = -b_1$ (notice that also $a_1$ could be zero) and

$$\begin{aligned} &a_{k,\ell} t^k s^\ell + a_{k-i,\ell+j} t^{k-i} s^{\ell+j} + a_{k-2i,\ell+2j} t^{k-2i} s^{\ell+2j} + \cdots + a_{k-ui,\ell+uj} t^{k-ui} s^{\ell+uj} \\ &= t^{k-ui} s^\ell a_1 (t^i - s^j)(a_2 t^i + b_2 s^j) \cdots (a_u t^i + b_u s^j). \end{aligned}$$

Let $u, v$ be integers such that $0 \le K - vi < i$ and $0 \le K - uj < j$, respectively. We deduce that (the second sum being unique, because of $\gcd(i,j) = 1$)

$$\begin{aligned} a(t,s) &= \sum_{k=0}^{K} \sum_{\ell=0}^{K} a_{k,\ell} t^k s^\ell \\ &= \sum_{k_1=0}^{i-1} \sum_{\ell_1=0}^{j-1} [\sum_{k_2=0}^{v} (a_{k_1+k_2 i,\ell_1} t^{k_1+k_2 i} s^{\ell_1} + a_{k_1+(k_2-1)i,\ell_1+j} t^{k_1+(k_2-1)i} s^{\ell_1+j} \end{aligned}$$

$$
\begin{aligned}
&+ a_{k_1+(k_2-2)i,\ell_1+2j}t^{k_1+(k_2-2)i}s^{\ell_1+2j} + \cdots \\
&+ a_{k_1+(k_2-k_2)i,\ell_1+k_2j}t^{k_1+(k_2-k_2)i}s^{\ell_1+k_2j}) \\
&+ \sum_{l_2=1}^{u}(a_{k_1+vi,\ell_1+\ell_2j}t^{k_1+vi}s^{\ell_1+\ell_2j} + a_{k_1+(v-1)i,\ell_1+(\ell_2+1)j}t^{k_1+(v-1)i}s^{\ell_1+(\ell_2+1)j} \\
&+ a_{k_1+(v-2)i,\ell_1+(\ell_2+2)j}t^{k_1+(v-2)i}s^{\ell_1+(\ell_2+2)j} + \cdots \\
&+ a_{k_1+(v-u+\ell_2)i,\ell_1+uj}t^{k_1+(v-u+\ell_2)i}s^{\ell_1+uj})] \\
= {} & (t^i - s^j)\sum b_{k,\ell}t^k s^\ell
\end{aligned}
$$

where if a subscript of $a_{k,\ell}$ is $> K$, then the $a_{k,\ell}$ in question is defined to be zero. ◻

PROPOSITION 3.2. *Let $N_p := p(p+3)/2$. The smallest $C$ satisfying the system of equations*

$$
\mathbf{q}_1^T C\mathbf{q}_i = 0,\ i = 1,2,\ldots,p, \quad \text{where} \quad \mathbf{q}_i = [1, q^i, q^{2i}, \ldots, q^{ni}]^T,\ i \in \mathbb{Z}^+,\ q \in \mathbb{R} \tag{3.3}
$$

*is $(N_p+1)\times(N_p+1)$. Furthermore, let $C_{N_p}$ be the smallest $C$ corresponding to $p$, then we can construct $C_{N_p}$ from $C_{N_{p-1}}$ according to the following prescription.*

$$
C_{N_1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{bmatrix},
$$

$$
\begin{aligned}
C_{N_p} = {} & \begin{bmatrix} O_{(p+1)\times N_{p-1}} & O_{(p+1)\times(p+1)} \\ C_{N_{p-1}} & O_{N_{p-1}\times(p+1)} \end{bmatrix} + \begin{bmatrix} C_{N_{p-1}} & O_{N_{p-1}\times(p+1)} \\ O_{(p+1)\times N_{p-1}} & O_{(p+1)\times(p+1)} \end{bmatrix} \\
& - \begin{bmatrix} O_{1\times 1} & O_{1\times N_{p-1}} & O_{1\times p} \\ O_{N_{p-1}\times 1} & C_{N_{p-1}} & O_{N_{p-1}\times p} \\ O_{p\times 1} & O_{p\times N_{p-1}} & O_{p\times p} \end{bmatrix} - \begin{bmatrix} O_{p\times p} & O_{p\times N_{p-1}} & O_{p\times 1} \\ O_{N_{p-1}\times p} & C_{N_{p-1}} & O_{N_{p-1}\times 1} \\ O_{1\times p} & O_{1\times N_{p-1}} & O_{1\times 1} \end{bmatrix},
\end{aligned} \tag{3.4}
$$

*where $p \geq 2$ and $O_{i\times j}$ is a zero matrix of order $i \times j$. Furthermore,*

$$
v_p(q) := \mathbf{1}^T C_{N_p}\mathbf{q}_1 = (1-q)^p \prod_{k=1}^{p}(1-q^k). \tag{3.5}
$$

*Proof.* We proceed by induction. By direct calculation, $\mathbf{q}_1^T C_{N_1}\mathbf{q}_1 = 0$. Assume ne xt that $\mathbf{q}_1^T C_{N_{p-1}}\mathbf{q}_i = 0$ for $i = 1,2,\ldots,p-1$. Then

$$
\begin{aligned}
\mathbf{q}_1^T C_{N_p}\mathbf{q}_i = {} & q^{p+1}\mathbf{q}_1^T C_{N_{p-1}}\mathbf{q}_i + \mathbf{q}_1^T C_{N_{p-1}}q^{i(p+1)}\mathbf{q}_i - q^1\mathbf{q}_1^T C_{N_{p-1}}q^i\mathbf{q}_i \\
& - q^p\mathbf{q}_1^T C_{N_{p-1}}q^{i\cdot p}\mathbf{q}_i = 0,
\end{aligned}
$$

for $i = 1,2,\ldots,p-1$. Furthermore,

$$
\begin{aligned}
\mathbf{q}_1^T C_{N_p}\mathbf{q}_p = {} & q^{p+1}\mathbf{q}_1^T C_{N_{p-1}}\mathbf{q}_p + \mathbf{q}_1^T C_{N_{p-1}}q^{p(p+1)}\mathbf{q}_p - q^1\mathbf{q}_1^T C_{N_{p-1}}q^p\mathbf{q}_p \\
& - q^p\mathbf{q}_1^T C_{N_{p-1}}q^{p\cdot p}\mathbf{q}_p = 0.
\end{aligned}
$$

Hence $\mathbf{q}_1^T C_{N_p}\mathbf{q}_i = 0$, for $i = 1,2,\ldots,p$ and any positive integer $p$.

To prove that $N_p$ is the smallest number such that there is at least one non-trivial $C$ satisfying (3.2), we write $\mathbf{S} = [1, s, \dots, s^n]^T$ and

$$b_i(q, s) = \mathbf{q}_1^T C_{N_i} \mathbf{S}, \quad i = 1, 2, \dots, p.$$

Then by our assumption

$$b_p(q, q^i) = \mathbf{q}_1^T C_{N_p} \mathbf{q}_i = 0, \quad i = 1, 2, \dots, p.$$

Hence, by the Lemma 3.1, $(s - q^i) | b_p(q, s), \quad i = 1, 2, \dots, p$. Note that $b_p(q, s)$ is a symmetric function of $q$ and $s$, thus

$$\prod_{k=1}^{p} [(s - q^k)(q - s^k)] | b_p(q, s).$$

On the other hand, by our recursion,

$$\begin{aligned}
b_1(q, s) &= \mathbf{q}_1^T C_{N_1} \mathbf{S} = q^2 - 2sq + s^2 = (q - s)(q - s), \\
b_p(q, s) &= \mathbf{q}_1^T C_{N_p} \mathbf{S} \\
&= q^{p+1} \mathbf{q}_1^T C_{N_{p-1}} \mathbf{S} + \mathbf{q}_1^T C_{N_{p-1}} s^{p+1} \mathbf{S} - q \mathbf{q}_1^T C_{N_{p-1}} s \mathbf{S} - q^p \mathbf{q}_1^T C_{N_{p-1}} s^p \mathbf{S} \\
&= (q^{p+1} + s^{p+1} - qs - q^p s^p) \mathbf{q}_1^T C_{N_{p-1}} \mathbf{S} \\
&= (q^p - s)(q - s^p) b_{p-1}(q, s) \\
&= \prod_{k=1}^{p} [(q^k - s)(q - s^k)] \qquad \text{(by induction)} \\
&= (-1)^p \prod_{k=1}^{p} [(s - q^k)(q - s^k)].
\end{aligned}$$

Thus $b_p(q, s)$ is the lowest degree $b(q, s)$ satisfying our assumption. The highest power of $s$ or $q$ in it, which corresponds to the smallest degree of matrix $C$, is

$$N_p = p + 1 + 2 + \cdots + p = \tfrac{1}{2} p(p + 3).$$

The next statement is obvious,

$$v_p(q) = b_p(q, 1) = \prod_{k=1}^{p} [(q^k - 1)(q - 1)] = (1 - q)^p \prod_{k=1}^{p} (1 - q^k).$$

The proof is complete. □

To ensure that

$$W(\varepsilon) := \varepsilon^{-1} \left( \mathbf{u}^T C \mathbf{u} - 2\hat{u} \mathbf{1}^T C \mathbf{u} \right) = O(\varepsilon^{p+1}),$$

it is necessary that (3.3) holds. However, for $p = 4$, (3.3) is not sufficient and we need one more equation,

$$\mathbf{q}_2^T C \mathbf{q}_3 = 0. \tag{3.6}$$

Generally, to guarantee $p \geq 4$,

$$\mathbf{q}_i^T C \mathbf{q}_j = 0, \qquad 2 \leq i + j \leq p + 1, \ \ 1 \leq i, j \leq p. \tag{3.7}$$

have to hold too. Thus, when $p$ gets larger, our matrix $C$ gets quite large. Fortunately, it is still possible to construct $C$ recursively.

THEOREM 3.3. *Let*

$$\mathbf{J}_p := \{k : 1 \le k \le [p/2];\ k \text{ and } p+1-k \text{ are relatively prime}\} =: \{k_1, k_2, \ldots, k_{J_p}\}$$

*where* $k_1 = 1$ *and* $J_p$ *is the number of elements in* $\mathbf{J}_p$. *Let*

$$\begin{aligned} M_1 &= 2,\\ M_p &= M_{p-1} + J_p(p+1) = \sum_{k=1}^{[p/2]} \sum_{\ell=k,\ \gcd(\ell,k)=1}^{p+1-k} (\ell+k), \quad p \ge 2,\end{aligned}$$

$$\begin{aligned} a_1(t,s) &= (t-s)^2,\\ a_p(t,s) &= a_{p-1}(t,s) \prod_{j=1}^{J_p} (t^{k_j} - s^{p+1-k_j})(t^{p+1-k_j} - s^{k_j})\\ &= \prod_{k=1}^{[p/2]} \prod_{\ell=k,\ \gcd(\ell,k)=1}^{p+1-k} (t^\ell - s^k)(t^k - s^\ell) := \sum_{i=0}^{M_p}\sum_{j=0}^{M_p} c_{i,j} t^i s^j \quad p \ge 2.\end{aligned}$$

*Then* $C = [c_{i,j}]_{i,j=0,1,\ldots,M_p}$ *is the smallest matrix satisfying* (3.7).

*We can construct* $C$ *by the following recursive scheme. Let*

$$C_{M_1} = \begin{bmatrix} 0 & 0 & 1\\ 0 & -2 & 0\\ 1 & 0 & 0 \end{bmatrix}.$$

*For* $p \ge 2$, *let* $M_{p,i} = M_{p-1} + i(p+1)$, $i = 0,1,\ldots,J_p$, $M_p := M_{p,J_p}$, *and set* $C_{M_{p,0}} := C_{M_{p-1}}$. *Then for* $i = 1,2,\ldots,J_p$,

$$\begin{aligned} C_{M_p,i} = &\begin{bmatrix} O_{(p+1)\times M_{p,i-1}} & O_{(p+1)\times(p+1)}\\ C_{M_{p,i-1}} & O_{M_{p,i-1}\times(p+1)} \end{bmatrix} + \begin{bmatrix} O_{M_{p,i-1}\times(p+1)} & C_{M_{p,i-1}}\\ O_{(p+1)\times(p+1)} & O_{(p+1)\times M_{p,i-1}} \end{bmatrix}\\ &- \begin{bmatrix} O_{k_i\times k_i} & O_{k_i\times M_{p,i-1}} & O_{k_i\times(p+1-k_i)}\\ O_{M_{p,i-1}\times k_i} & C_{M_{p,i-1}} & O_{M_{p,i-1}\times(p+1-k_i)}\\ O_{(p+1-k_i)\times k_i} & O_{(p+1-k_i)\times M_{p,i-1}} & O_{(p+1-k_i)\times(p+1-k_i)} \end{bmatrix}\\ &- \begin{bmatrix} O_{(p+1-k_i)\times(p+1-k_i)} & O_{(p+1-k_i)\times M_{p,i-1}} & O_{(p+1-k_i)\times k_i}\\ O_{M_{p,i-1}\times(p+1-k_i)} & C_{M_{p,i-1}} & O_{M_{p,i-1}\times k_i}\\ O_{k_i\times(p+1-k_i)} & O_{k_i\times M_{p,i-1}} & O_{k_i\times k_i} \end{bmatrix}.\end{aligned}$$

*Furthermore,*

$$v_p(q) = v_{p-1}(q)\prod_{i=1}^{J_p}(1-q^{k_i})(1-q^{p+1-k_i}) = \prod_{k=1}^{[p/2]} \prod_{\ell=k,\ \gcd(\ell,k)=1}^{p+1-k} (1-q^\ell)(1-q^k).$$

*Proof.* The proof is straightforward. Suppose $1 \le i \le p$, $1 \le j \le p$, $2 \le i+j \le p+1$. If $\gcd(i,j) = 1$, then the factor $(t^i - s^j)$ is a divisor of $a_p(t,s)$, thus

$$\mathbf{q}_i^T C \mathbf{q}_j = a_p(q^j, q^i) = 0.$$

On the other hand, if $\gcd(i,j) = \sigma \geq 2$ then $i = k\sigma$, $j = l\sigma$ and $1 \leq k, l < p$, $2 \leq k + l < p + 1$, and we still have

$$\mathbf{q}_i^T C \mathbf{q}_j = \mathbf{q}_k^T C \mathbf{q}_l = a_p(q^l, q^k) = 0.$$

Hence (3.7) is satisfied.

Furthermore, our Lemma 3.1 guarantees that $C = C_{M_p}$ is the smallest matrix satisfying (3.7).

We can get our recursive scheme for $C$ from the recurrence formula for $a_p(t,s)$. The last statement is obvious considering $v_p(t) = a_p(1,t)$. □

As for $n \geq M_p + 1$, we can find more than one matrix $C$ that satisfies (3.7).

PROPOSITION 3.4. *Let $\mathcal{C}_p^n$ be the space of matrices satisfying* (3.7). *The dimension of $\mathcal{C}_p^n$ is $(n - M_p + 1)(n - M_p + 2)/2$, where $M_p$ has been defined in Theorem 3.3. Furthermore, for fixed p and $M_p$, we can construct a base of $\mathcal{C}_p^n$ in the following way. There is one element, up to a constant multiple, in $\mathcal{C}_p^{M_p}$ as constructed before, which we write as $C_{M_p}$. All the three elements in $\mathcal{C}_p^{M_p+1}$ can be written in the form*

$$C^{(1)}_{M_p+1} = \begin{bmatrix} O_{1\times M_p+1} & 0 \\ C_{M_p} & O_{M_p+1\times 1} \end{bmatrix} + \begin{bmatrix} O_{M_p+1\times 1} & C_{M_p} \\ 0 & O_{1\times M_p+1} \end{bmatrix},$$
$$C^{(2)}_{M_p+1} = \begin{bmatrix} 0 & O_{1\times M_p+1} \\ O_{M_p+1\times 1} & C_{M_p} \end{bmatrix}, \qquad C^{(3)}_{M_p+1} = \begin{bmatrix} C_{M_p} & O_{M_p+1\times 1} \\ O_{1\times M_p+1} & 0 \end{bmatrix}.$$

*Furthermore,*

$$v^{(0)}_{p,p+1}(q) = \mathbf{1}^T C^{(1)}_{M_p+1}\mathbf{q}_1 = \mathbf{1}^T C_{M_p}\mathbf{q}_1 + \mathbf{1}^T C_{M_p} q\mathbf{q}_1 = (1+q)v_p(q);$$
$$v^{(1)}_{p,p+1}(q) = \mathbf{1}^T C^{(2)}_{M_p+1}\mathbf{q}_1 = \mathbf{1}^T C_{M_p} q\mathbf{q}_1 = q v_p(q);$$
$$v^{(2)}_{p,p+1}(q) = \mathbf{1}^T C^{(3)}_{M_p+1}\mathbf{q}_1 = \mathbf{1}^T C_{M_p}\mathbf{q}_1 = v_p(q).$$

*For general $n \geq M_p + 1$ we can construct a base of $\mathcal{C}_p^n$ by recursion. Let $L = n - M_p$. First we construct one matrix of $(M_p + k + 1) \times (M_p + k + 1)$ by enlarging $C_{M_p}$ in the following fashion,*

$$C^{(0)}_{M_p,0} = C_{M_p};$$
$$C^{(0)}_{M_p+k,k} = \begin{bmatrix} O_{1\times M_p+k} & 0 \\ C^{(0)}_{M_p+k-1,k-1} & O_{M_p+k\times 1} \end{bmatrix} + \begin{bmatrix} O_{M_p+k\times 1} & C^{(0)}_{M_p+k-1,k-1} \\ 0 & O_{1\times M_p+k} \end{bmatrix}$$

*for $k = 1, 2, \ldots, L$. We augment each to a matrix of size $(M_p + L + 1) \times (M_p + L + 1)$ by adding $L - k$ zero rows and columns as follows, whereby they are all linearly independent.*

$$C^{(i)}_{M_p+L,k} = \begin{bmatrix} O_{i\times i} & O_{i\times M_p+k+1} & O_{i\times L-k-i} \\ O_{M_p+k+1\times i} & C^{(0)}_{M_p+k,k} & O_{M_p+k+1\times L-k-i} \\ O_{L-k-i\times i} & O_{L-k-i\times M_p+k+1} & O_{L-k-i\times L-k-i} \end{bmatrix},$$

$$i = 0, 1, \ldots, L - k,\; k = 0, 1, 2, \ldots, L.$$

*Correspondingly, $v^{(i,k)}_{p,n}(q) := \mathbf{1}^T C^{(i)}_{M_p+L,k}\mathbf{q}_1 = q^i(1+q)^k v_p(q)$, $i = 0, 1, 2, \ldots, L - k$, $k = 0, 1, 2, \ldots, L$.*

*Proof.* Let $C \in \mathcal{C}_p^n$ and

$$b(q,s) = \mathbf{q}_1^T C \mathbf{S}.$$

Then, according to the lemma,

$$\prod_{k=1}^{[p/2]} \prod_{\ell=k,\gcd(\ell,k)=1}^{p+1-k} (q^\ell - s^k)(q^k - s^\ell) | b(q,s).$$

Hence

$$b(q,s) = \prod_{k=1}^{[p/2]} \prod_{\ell=k,\ \gcd(\ell,k)=1}^{p+1-k} (q^\ell - s^k)(q^k - s^\ell) r(q,s),$$

where $r(s,q)$ is a polynomial of degree $n - M_p$ in two variables. There are $(n - M_p + 1)(n - M_p + 2)/2$ unknowns for $r(t,s)$, that is the upper bound for the dimension of $\mathcal{C}_p^n$. It is obvious that our construction gives $(n - M_p + 1)(n - M_p + 2)/2$ linearly independent matrices. Thus our assertions are true. ☐

THEOREM 3.5. *If we choose* $C \in \mathcal{C}_p^{M_p}$ *as in Theorem 3.3 or* $C \in \mathcal{C}_p^n$ *as in Proposition 3.4, then the quadratic model*

$$G(z) = \frac{\mathbf{u}^T C \mathbf{u}}{2\mathbf{1}^T C \mathbf{u}}$$

*has at least order* $p$*, unless* $\beta_1$ *is a root of unity.*

**Examples.** For $p = 2$, $n = 5$, $C$ is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & -2 & 0 \\ 0 & 0 & 2 & 1 & -1 & 0 \\ 0 & -1 & 1 & 2 & 0 & 0 \\ 0 & -2 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.8}$$

For $p = 2$, $n = 6$, all the $C$s are

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & -2 & 0 \\ 0 & 0 & 0 & 2 & 1 & -1 & 0 \\ 0 & 0 & -1 & 1 & 2 & 0 & 0 \\ 0 & 0 & -2 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & -2 & 0 & 0 \\ 0 & 0 & 2 & 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 2 & 0 & 0 & 0 \\ 0 & -2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

For $p = 2$, $n = 7$, the $C$s are

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & -2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & -2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

$$\begin{bmatrix} 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&-1&-2&0\\ 0&0&0&0&2&1&-1&0\\ 0&0&0&-1&1&2&0&0\\ 0&0&0&-2&-1&0&0&0\\ 0&0&1&0&0&0&0&0 \end{bmatrix};\quad \begin{bmatrix} 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&-1&-1&0\\ 0&0&0&0&1&-1&-1&0\\ 0&0&0&1&2&1&0&0\\ 0&0&-1&-1&1&0&0&0\\ 0&0&-1&-1&0&0&0&0\\ 0&1&0&0&0&0&0&0 \end{bmatrix};\quad \begin{bmatrix} 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&-1&0&0\\ 0&0&0&0&0&-2&-1&0\\ 0&0&0&2&1&0&0&0\\ 0&0&0&1&2&0&0&0\\ 0&-1&-2&0&0&0&0&0\\ 0&0&-1&0&0&0&0&0\\ 1&0&0&0&0&0&0&0 \end{bmatrix}.$$

Finally, for $p = 3$, $n = 9$, the matrix $C$ is

$$\begin{bmatrix} 0&0&0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&-1&-1&-2&0\\ 0&0&0&0&1&2&2&1&-1&0\\ 0&0&0&-2&-1&0&1&2&-1&0\\ 0&0&1&-1&-2&-1&0&2&0&0\\ 0&0&2&0&-1&-2&-1&1&0&0\\ 0&-1&2&1&0&-1&-2&0&0&0\\ 0&-1&1&2&2&1&0&0&0&0\\ 0&-2&-1&-1&0&0&0&0&0&0\\ 1&0&0&0&0&0&0&0&0&0 \end{bmatrix}.$$

Let us study next the attractivity of $G_N(z)$.

THEOREM 3.6. *Given $p > 0$. Let*

$$\begin{aligned} \mathbf{J}_p &:= \{k : 1 \le k \le [p/2];\ k \text{ and } p+1-k \text{ are relatively prime }\} = \{k_1, k_2, \dots, k_{J_p}\},\\ S_p &= k_1 + \cdots + k_{J_p},\\ K_p &= S_1 + \cdots + S_p, \end{aligned}$$

*where $k_1 = 1$ and $J_p$ is the number of elements in $\mathbf{J}_p$. If $f(z)$ is super-attractive of degree $s \ge 1$ at the fixed point $\hat{z}$, then $G_{M_p}(z)$ is super-attractive there of the degree $2(s+1)^{K_p} - 2$.*

*Proof.* Without loss of generality we suppose that $\hat{z} = 0$. If

$$f(z) = \alpha_0 z^{s+1} + O(z^{s+2}),$$

then by induction we have

$$f^{\circ m}(z) = \alpha_m z^{(s+1)^m} + O(z^{(s+1)^m+1}).$$

Note that

$$G_{M_p}(z) = \frac{\mathbf{u}^T C_{M_p} \mathbf{u}}{2\mathbf{1}^T C_{M_p} \mathbf{u}} = \frac{\sum_{i,j} c_{i,j} f^{\circ i} f^{\circ j}}{2\sum_{i,j} c_{i,j} f^{\circ i}}.$$

By the construction of $C_{M_p}$ we can see that $c_{i,j} = 0$ for $0 \le i + j \le 2(K_p) - 1$ and $c_{K_p,K_p} \ne 0$. Among the terms of $c_{i,j} f^{\circ i} f^{\circ j}$ the smallest power of $z$ is in the form $c_{i,j} z^{(s+1)^i+(s+1)^j}$ when $s \ge 1$, for $2K_p \le i + j \le 2M_p$ . Furthermore, $\sum_{j=0}^{M_p} c_{0,j} = c_{0,M_p} = 1$. Thus,

$$\begin{aligned} G_{M_p}(z) &= \frac{c_{K_p,K_p} \alpha_{K_p}^2 z^{2(s+1)^{K_p}} + O(z^{2(s+1)^{K_p}+1})}{c_{0,M_p} z + O(z^2)}\\ &= c_{K_p,K_p} \alpha_{K_p}^2 z^{2(s+1)^{K_p}-1} + O(z^{2(s+1)^{K_p}}). \qquad \square \end{aligned}$$

Corollary. *Let*

$$G_{n,p}(z) = \frac{\mathbf{u}^T C_{n,0}^{(n-M_p)} \mathbf{u}}{2\mathbf{1}^T C_{n,0}^{(n-M_p)} \mathbf{u}}$$

*with* $C_{n,0}^{(n-M_p)}$ *as in Proposition 3.4, then if* $f(z)$ *is super-attractive of degree* $s \geq 1$ *at the fixed point* $\hat{z}$ *then* $G_{n,p}(z)$ *is super-attractive there of degree* $2(s+1)^{n-M_p+K_p} - (s+1)^{n-M_p} - 1$.

**4. The fixed point at $\infty$.** Let $f$ be analytic for $|z| \geq 1$ and $\infty \in F_f$. Assume that

$$f(z) = \gamma_0 z + \sum_{k=1}^{\infty} \frac{1}{k!} \gamma_k z^{-k+1}.$$

For the time being, we assume that $\gamma_0$ is neither zero nor a root of unity. Then the Taylor expansion of $f^{\circ m}$ about $\infty$ is

$$f^{\circ m}(z) = \gamma_0^m z + \sum_{k=1}^{\infty} \frac{1}{k!} \epsilon_k^{(m)} z^{-k+1} \qquad m = 1, 2, \ldots.$$

According to Iserles [2], there exist numbers $D_{k,l}$, $l = 0, 1, \ldots, k$, $k = 0, 1, \ldots$, dependent on $\{\gamma_k\}_{k=1}^{\infty}$ but not on $m$, such that

$$\epsilon_k^{(m+1)} = \sum_{l=0}^{k} D_{k,l} \epsilon_l^{(m)}.$$

Moreover, $D_{k,k} = \gamma_0^{1-k}$, $k = 0, 1, \ldots$. Let

$$B_{1,1} = 1,$$

$$B_{k,l} = \frac{1}{\gamma_0^l - \gamma_0^{1-k}} \sum_{j=1-l}^{k-1} D_{k,j} B_{j,l}, \qquad l = 2-k, 3-k, \ldots, 1, \quad k = 1, 2, \ldots,$$

$$B_{k,1-k} = \gamma_0^{k-1} (\gamma_k - \sum_{l=2-k}^{1} B_{k,l} \gamma_0^l), \qquad k = 1, 2, \ldots.$$

Then

$$\epsilon_k^{(m)} = \sum_{l=1-k}^{1} B_{k,l} \gamma_0^{lm} \qquad k, m = 0, 1, \ldots.$$

Therefore if $\mathbf{p}_1^T C \mathbf{p}_1 = 0$, $\mathbf{1}^T C \mathbf{p}_1 \neq 0$, then

$$G(z) = \frac{\mathbf{u}^T C \mathbf{u}}{2\mathbf{1}^T C \mathbf{u}} = \frac{\mathbf{p}_1^T C \mathbf{p}_1 z^2 + 2B_{1,0} \mathbf{1}^T C \mathbf{p}_1 z + O(1)}{2\mathbf{1}^T C \mathbf{p}_1 z + \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{l=1-k}^{1} B_{k,l} \mathbf{1}^T C \mathbf{p}_l z^{-k+1}} = B_{1,0} + O(z^{-1}).$$

In the case when $\infty$ is a super-attractive fixed point of $f$ of degree $s$, we write

$$f(z) = \gamma_0 z^s + O(z^{s-1}).$$

Then

$$f^{\circ m}(z) = \gamma_0^{(m)} z^{s^m} + O(z^{s^m - 1}).$$

Note the highest power among $z^{s^i + s^j}$, for $1 \leq i, j \leq M_p$, must be located at $i = M_p$ or $j = M_p$ and the fact that in the matrix $C_{M_p}$, $c_{M_p,i} = c_{i,M_p} = 0$ for $1 \leq i \leq M_p$. Therefore we get

$$\begin{aligned} G_{M_p}(z) &= \frac{\mathbf{u}^T C_{M_p} \mathbf{u}}{2\mathbf{1}^T C_{M_p} \mathbf{u}} = \frac{\sum_{i,j} c_{i,j} f^{\circ i}(z) f^{\circ j}(z)}{2 \sum_{i,j} c_{i,j} f^{\circ i}(z)} \\ &= \frac{2c_{0,M_p} \gamma_0^{(M_p)} z^{s^{M_p}+1} + O(z^{s^{M_p}})}{2c_{0,M_p} \gamma_0^{(M_p)} z^{s^{M_p}} + O(z^{s^{M_p}-1})} = z + O(1). \end{aligned}$$

We have thus deduced the following result.

Proposition 4.1. *If $\infty$ is a fixed point of $f$ which is neither neutral with $\gamma_0$ a root of unity nor super-attractive, then $\infty$ is not a fixed point of $G_{M_p}(z)$ for*

$$G_{M_p}(z) = \frac{\mathbf{u}^T C_{M_p} \mathbf{u}}{2\mathbf{1}^T C_{M_p} \mathbf{u}}$$

Proposition 4.2. *If $\infty$ is a super-attractive fixed point of $f$ with $s \geq 2$, then $\infty$ is a fixed point with degree 1 of $G_{M_p}(z)$ for*

$$G_{M_p}(z) = \frac{\mathbf{u}^T C_{M_p} \mathbf{u}}{2\mathbf{1}^T C_{M_p} \mathbf{u}}$$

*For a general $n \geq M_p + 1$, we choose $C_{n,0}^{(n-M_p)} \in \mathcal{C}_p^n$ and $p \geq 1$, since*

$$\begin{aligned} G_{n,p}(z) &= \frac{\mathbf{u}^T C_{n,0}^{(n-M_p)} \mathbf{u}}{2\mathbf{1}^T C_{n,0}^{(n-M_p)} \mathbf{u}} = \frac{\sum_{i,j} c_{i,j} f^{\circ i}(z) f^{\circ j}(z)}{2 \sum_{i,j} c_{i,j} f^{\circ i}(z)} \\ &= \frac{2c_{n-M_p,n} \gamma_0^{(n-M_p)} \gamma_0^{(n)} z^{s^{n-M_p}+s^n} + O(z^{s^{n-M_p}+s^n-1})}{2c_{n-M_p,n} \gamma_0^{(n)} z^{s^n} + O(z^{s^n-1})} \\ &= \gamma_0^{(n-M_p)} z^{s^{n-M_p}} + O(z^{s^{n-M_p}-1}), \end{aligned}$$

*the point $\infty$ is still a super-attractive fixed point up to degree $s^{n-M_p} - 1$.*

Proposition 4.3. *If $\infty$ is a fixed point of $f$ which is neither neutral with $\gamma_0$ a root of unity nor supper-attractive, then $\infty$ is not a fixed point of $G(z)$ for*

$$G(z) = \frac{\mathbf{u}^T C \mathbf{u}}{2\mathbf{1}^T C \mathbf{u}} \quad \textit{where} \quad C \in \mathcal{C}_p^n \quad \textit{and} \quad p \geq 1; \;\; n \geq M_p + 1.$$

*If $\infty$ is a super-attractive fixed point of $f$ with degree $s \geq 2$ then $\infty$ is a super-attractive fixed point of degree up to $s^{n-M_p} - 1$ for the same $G(z)$ as above.*

**5. The behaviour for $\beta_1 = 1$.** Supposing that

$$f(z) = z + \sum_{\ell=2}^{\infty} \tfrac{1}{\ell!}\beta_\ell z^\ell$$

where, again, $f^{\circ m}$ is the $m$th iterant of $f$, it follows from [1] that there exist constants $r_{\ell,j}$ such that

$$f^{\circ m}(z) = z + \sum_{\ell=2}^{\infty} \left( \sum_{j=1}^{\ell-1} r_{\ell,j} m^j \right) z^\ell, \qquad m \in \mathbb{Z}^+, \tag{5.1}$$

where

$$r_{\ell,\ell-1} = \left(\frac{\beta_2}{2}\right)^{\ell-1}, \qquad \ell = 2, 3, \ldots.$$

Suppose that $C \in \mathcal{C}_1^n$. Differentiating $\mathbf{q}_1^T C \mathbf{q}_1 \equiv 0$with respect to $q$ we obtain

$$\mathbf{q}_1^T C \mathbf{q}_1' \equiv 0,$$

hence, letting $q = 1$,

$$\mathbf{1}^T C \mathbf{k}_1 = 0, \tag{5.2}$$

where

$$\mathbf{k}_\ell^T := \left[\; 0^\ell \; 1^\ell \; 2^\ell \; \cdots \; n^\ell \;\right], \qquad \ell \in \mathbb{Z}^+.$$

Differentiating again, we have

$$\mathbf{q}_1^T C \mathbf{q}_1'' + \mathbf{q}_1'^T C \mathbf{q}_1' = 0,$$

hence, because of (5.2),

$$\mathbf{1}^T C \mathbf{k}_2 + \mathbf{k}_1^T C \mathbf{k}_1 = 0. \tag{5.3}$$

Because of (5.1), we have

$$\mathbf{u} = \mathbf{1} z + \sum_{\ell=2}^{\infty} \left( \sum_{j=1}^{\ell-1} r_{\ell,j} \mathbf{k}_j \right) z^\ell.$$

Thus, taking (5.2) into account,

$$\mathbf{1}^T C \mathbf{u} = \sum_{\ell=3}^{\infty} \left( \sum_{j=2}^{\ell-1} r_{\ell,j} \mathbf{1}^T C \mathbf{k}_j \right) z^\ell,$$

$$\mathbf{u}^T C \mathbf{u} = 2 \sum_{\ell=3}^{\infty} \left( \sum_{j=2}^{\ell-1} r_{\ell,j} \mathbf{1}^T C \mathbf{k}_j \right) z^{\ell+1} + \sum_{\ell_1=2}^{\infty} \sum_{\ell_2=2}^{\infty} \sum_{j_1=1}^{\ell_1-1} \sum_{j_2=1}^{\ell_2-1} r_{\ell_1,j_1} r_{\ell_2,j_2} \mathbf{k}_{j_1}^T C \mathbf{k}_{j_2} z^{\ell_1+\ell_2}.$$

Recall that the function that we are iterating is

$$G_n(z) := \tfrac{1}{2} \frac{\mathbf{u}^T C \mathbf{u}}{\mathbf{1}^T C \mathbf{u}}.$$

In our case

$$G_n(z) = z\left(1 + \frac{r_{2,1}^2 \mathbf{k}_1^T C\mathbf{k}_1 + O(z)}{2r_{3,2}\mathbf{1}^T C\mathbf{k}_2 + O(z)}\right).$$

Thus, unless $\mathbf{k}_1^T C\mathbf{k}_1 = 0$, (5.3) yields

$$G_n(z) = \tfrac{1}{2}z + O(z^2).$$

In other words, the fixed point is merely attractive (not super-attractive!) and $G_1'(0) = \frac{1}{2}$. This is identical to the standard Steffensen method [2].

The remaining case is

$$\mathbf{k}_1^T C\mathbf{k}_1 = \mathbf{1}^T C\mathbf{k}_2 = 0. \tag{5.4}$$

Considering the third derivative of $\mathbf{q}_1^T C\mathbf{q}_1 \equiv 0$, we readily affirm (taking into account (5.2)–(5.4)) that

$$\mathbf{1}^T C\mathbf{k}_3 + 3\mathbf{k}_1^T C\mathbf{k}_2 = 0. \tag{5.5}$$

Now

$$G_n(z) = z\left(1 + \frac{r_{2,1}r_{3,2}\mathbf{k}_1^T C\mathbf{k}_2 + O(z)}{r_{4,3}\mathbf{1}^T C\mathbf{k}_3 + O(z)}\right).$$

Thus, because of (5.5) and unless $\mathbf{k}_1^T C\mathbf{k}_2 = 0$, we have

$$G_n(z) = \tfrac{2}{3}z + O(z^2)$$

and the situation is actually *worse* than in the previous case!

So, let us proceed a step further, replacing (5.5) with

$$\mathbf{1}^T C\mathbf{k}_3 = \mathbf{k}_1^T C\mathbf{k}_2 = 0. \tag{5.6}$$

Another differentiation of $\mathbf{q}_1^T C\mathbf{q}_1 \equiv 0$ gives

$$\mathbf{1}^T C\mathbf{k}_4 + 4\mathbf{k}_1^T C\mathbf{k}_3 + 3\mathbf{k}_2^T C\mathbf{k}_2 = 0. \tag{5.7}$$

In the present case

$$\begin{aligned} G_n(z) &= z\left(1 + \frac{2r_{2,1}r_{4,3}\mathbf{k}_1^T C\mathbf{k}_3 + r_{3,2}^2\mathbf{k}_2^T C\mathbf{k}_2 + O(z)}{2r_{5,4}\mathbf{1}^T C\mathbf{k}_4}\right) \\ &= z\left(1 + \frac{2\mathbf{k}_1^T C\mathbf{k}_3 + \mathbf{k}_2^T C\mathbf{k}_2 + O(z)}{\mathbf{1}^T C\mathbf{k}_4 + O(z)}\right). \end{aligned}$$

Thus, provided that $\mathbf{1}^T C\mathbf{k}_4 \neq 0$, $G_n(z) = O(z^2)$ if and only if

$$\mathbf{1}^T C\mathbf{k}_4 + 2\mathbf{k}_1^T C\mathbf{k}_3 + \mathbf{k}_2^T C\mathbf{k}_2 = 0.$$

Substituting the value of $\mathbf{1}^T C\mathbf{k}_4$ from (5.7) yields the conditions

$$\mathbf{1}^T C\mathbf{k}_4 \neq 0, \qquad \mathbf{k}_1^T C\mathbf{k}_3 + \mathbf{k}_2^T C\mathbf{k}_2 = 0. \tag{5.8}$$

It is instructive to check (5.8) in $\mathcal{C}_2^n$ for $n \in \{5, \ldots, 7\}$. For $n = 5$ we have a one-dimensional space spanned by (3.8), and the latter gives

$$\mathbf{k}_1^T C\mathbf{k}_3 + \mathbf{k}_2^T C\mathbf{k}_2 = 2.$$

Likewise, for $n = 6$ we have

$$\mathbf{k}_1^T C\mathbf{k}_3 + \mathbf{k}_2^T C\mathbf{k}_2 = 4.$$

However, in the case $n = 7$ we have a two-dimensional space and

$$\begin{aligned} \mathbf{k}_1^T P_1\mathbf{k}_3 + \mathbf{k}_2^T P_1\mathbf{k}_2 &= 8, \\ \mathbf{k}_1^T P_2\mathbf{k}_3 + \mathbf{k}_2^T P_2\mathbf{k}_2 &= 2. \end{aligned}$$

Hence, the only possible choice of $C$ (up to a nonzero multiplicative constant) which is consistent with the second condition in (5.8) is

$$P := P_1 - 4P_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -4 & 0 \\ 0 & 0 & 0 & 0 & 4 & 6 & -1 & 0 \\ 0 & 0 & 0 & -6 & -3 & 4 & 0 & 0 \\ 0 & 0 & 4 & -3 & -6 & 0 & 0 & 0 \\ 0 & -1 & 6 & 4 & 0 & 0 & 0 & 0 \\ 0 & -4 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Unfortunately, the first condition of (5.8) is violated, since $\mathbf{1}^T C\mathbf{k}_4 = 0$. This, in fact, is predictable – the function $v$ for $P$ is of the form

$$v(t) = v_1(t) - 4v_2(t) = (1-t)^6(1+t).$$

Thus, $\mathrm{d}^\ell v(1)/\,\mathrm{d}t^\ell = 0$, $\ell = 0, 1, \ldots, 5$. But

$$\begin{aligned} v(1) &= \mathbf{1}^T C\mathbf{1}, \\ v'(1) &= \mathbf{1}^T C\mathbf{k}_1, \\ v''(1) &= \mathbf{1}^T C(\mathbf{k}_2 - \mathbf{k}_1), \\ v'''(1) &= \mathbf{1}^T C(\mathbf{k}_3 - 3\mathbf{k}_2 + 2\mathbf{k}_1) \end{aligned}$$

and so on. We conclude that $\mathbf{1}^T C\mathbf{k}_\ell = 0$, $\ell = 0, \ldots, 5$. The general result can be phrased as a theorem.

Theorem 5.1. *Suppose that*

$$f(z) = z + \sum_{\ell=2}^{\infty} \frac{1}{\ell!}\beta_\ell z^\ell$$

$$G(z) = \frac{\mathbf{u}^T C\mathbf{u}}{2\mathbf{1}^T C\mathbf{u}} \qquad \textit{where} \quad C \in \mathcal{C}_p^{M_p} \quad \textit{and} \quad p \ge 1.$$

*Then*

$$G(z) = \alpha_p z + O(z^2) \qquad \textit{for} \quad p \ge 1$$

*where $\alpha_p \neq 0$ for $p = 1, 2, \ldots, 5$. Furthermore, letting $L_p = J_1 + \cdots + J_p$, it is true that $\sum_{j_1+j_2=2L_p} \mathbf{k}_{j_1}^T C_{M_p} \mathbf{k}_{j_2} \neq 0$ for $p \geq 6$ implies $\alpha_p \neq 0$.*

*Proof.* Since $p \geq 1$, we have

$$\mathbf{1}^T C\mathbf{u} = \sum_{\ell=3}^{\infty} \left( \sum_{j=2}^{\ell-1} r_{\ell,j} \mathbf{1}^T C\mathbf{k}_j \right) z^{\ell},$$

$$\mathbf{u}^T C\mathbf{u} = 2\sum_{\ell=3}^{\infty} \left( \sum_{j=2}^{\ell-1} r_{\ell,j} \mathbf{1}^T C\mathbf{k}_j \right) z^{\ell+1} + \sum_{\ell_1=2}^{\infty} \sum_{\ell_2=2}^{\infty} \sum_{j_1=1}^{\ell_1-1} \sum_{j_2=1}^{\ell_2-1} r_{\ell_1,j_1} r_{\ell_2,j_2} \mathbf{k}_{j_1}^T C\mathbf{k}_{j_2} z^{\ell_1+\ell_2}.$$

We will prove that, letting $\mathbf{k}_0 := \mathbf{1}$,

$$\mathbf{k}_i^T C_{M_p} \mathbf{k}_j = 0, \quad \text{for} \quad 0 \leq i + j \leq 2L_p - 1. \tag{5.9}$$

Let

$$a_1(t,s) = (t-s)(t-s);$$

$$a_p(t,s) = a_{p-1}(t,s) \prod_{k=1,\ \gcd(k,p+1-k)=1}^{[p/2]} (t^k - s^{p+1-k})(t^{p+1-k} - s^k)$$

$$= \prod_{k=1}^{[p/2]} \prod_{i=k,\gcd(i,k)=1}^{p+1-k} (t^i - s^k)(t^k - s^i).$$

Then, since $a_p(t,s)$ has $2L_p$ factors of form $(t^k - s^l)$ and these factors are zero when $t = 1$ and $s = 1$,

$$\mathbf{k}_i^T C_{M_p} \mathbf{k}_j = \frac{\partial}{\partial t}\left( t\frac{\partial}{\partial t}\left( t \cdots \frac{\partial}{\partial t}\left( t\frac{\partial}{\partial s}\left( s\frac{\partial}{\partial s}\left( s \cdots \frac{\partial}{\partial s}(a_p(t,s)) \cdots \right)\right)\right) \cdots \right)\right)\Bigg|_{t=1,\ s=1}$$

$$= \left( \frac{\partial^{i+j}}{\partial^i t \partial^j s} a_p(t,s) + \sum_{k+l<i+j, 0\leq k\leq i, 0\leq l\leq j} \sigma_{k,l} \frac{\partial^{k+l}}{\partial^k t \partial^l s} a_p(t,s) \right)\Bigg|_{t=1,\ s=1}$$

$$= 0, \quad \text{for} \quad 0 \leq i + j \leq 2L_p - 1,$$

Furthermore,

$$\mathbf{1}^T C_{M_p} \mathbf{k}_{2L_p} = \left[ \frac{\partial^{2L_p}}{\partial^{2L_p} s} a_p(t,s) + \sum_{k+l<2L_p} \sigma_{k,l} \frac{\partial^{k+l}}{\partial^k t \partial^l s} a_p(t,s) \right]\Bigg|_{t=1,\ s=1}$$

$$= (2L_p)! \prod_{k=1}^{[p/2]} \prod_{i=k,\ \gcd(i,k)=1}^{p+1-k} ik \neq 0.$$

Consequently,

$$\mathbf{1}^T C_{M_p} \mathbf{u} = \sum_{l=2L_p+1}^{\infty} \left( \sum_{j=2L_p}^{l-1} r_{l,j} \mathbf{1}^T C_{M_p} \mathbf{k}_j \right) z^l$$

$$= \left(\frac{\beta_2}{2}\right)^{2L_p} \mathbf{1}^T C_{M_p} \mathbf{k}_{2L_p} z^{2L_p+1} + O(z^{2L_p+2})$$

$$\mathbf{u}^T C_{M_p} \mathbf{u} = 2 \sum_{l=2L_p+1}^{\infty} \left( \sum_{j=2L_p}^{l-1} r_{l,j} \mathbf{1}^T C_{M_p} \mathbf{k}_j \right) z^{l+1}$$

$$+ \sum_{l_1+l_2>2L_p+1}^{\infty} \sum_{j_1=1}^{l_1-1} \sum_{j_2=2L_p-j_1}^{l_2-1} r_{l_1,j_1} r_{l_2,j_2} \mathbf{k}_{j_1}^T C_{M_p} \mathbf{k}_{j_2} z^{l_1+l_2}$$

$$= 2r_{2L_p+1,2L_p} \mathbf{1}^T C_{M_p} \mathbf{k}_{2p} z^{2L_p+2}$$

$$+ \sum_{j_1,j_2>0, j_1+j_2=2L_p} r_{j_1+1,j_1} r_{j_2+1,j_2} \mathbf{k}_{j_1}^T C_{M_p} \mathbf{k}_{j_2} z^{2L_p+2} + O(z^{2L_p+3})$$

$$= \left(\frac{\beta_2}{2}\right)^{2L_p} \sum_{j_1+j_2=2L_p} \mathbf{k}_{j_1}^T C_{M_p} \mathbf{k}_{j_2} z^{2L_p+2} + O(z^{2L_p+3}),$$

where $\mathbf{k}_0 = \mathbf{1}$. So we conclude

$$G(z) = \alpha_p z + O(z^2) \quad \text{for} \quad p \geq 1,$$

where $\alpha_p \neq 0$ if $\sum_{j_1+j_2=2L_p} \mathbf{k}_{j_1}^T C_{M_p} \mathbf{k}_{j_2} \neq 0$. By direct calculation we can easily deduce that $\alpha_p \neq 0$ for $p = 1, 2, 3, 4, 5$. This completes the proof. □

## REFERENCES


[1] C. Brezinski and M. Redivo Zaglia, *Extrapolation methods. Theory and Practice,* North-Holland, Amsterdam, 1993.

[2] A. Iserles, *Complex dynamics of convergence acceleration*, IMA J. of Numer. Anal., 11 (1991), pp. 205–240.